

OIPE

DATE: 10/26/2001
TIME: 13:23:11

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/829,631A

Input Set : A:\NIH047.1CP1C1.TXT
Output Set: N:\CRF3\10262001\I829631A.raw

4 <110> APPLICANT: Sibley, David R.
5 Monsma, Frederick J.
6 Hamblin, Mark
9 <120> TITLE OF INVENTION: The ST-B17 Serotonin Receptor
12 <130> FILE REFERENCE: NIH047.1CP1C1
14 <140> CURRENT APPLICATION NUMBER: US 09/829,631A
15 <141> CURRENT FILING DATE: 2001-04-10
17 <150> PRIOR APPLICATION NUMBER: US 08/428,242
18 <151> PRIOR FILING DATE: 1995-09-18
20 <160> NUMBER OF SEQ ID NOS: 13
22 <170> SOFTWARE: FastSEQ for Windows Version 4.0
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 33
26 <212> TYPE: DNA
27 <213> ORGANISM: Artificial Sequence
29 <220> FEATURE:
30 <223> OTHER INFORMATION: primer
33 <400> SEQUENCE: 1
34 atcqaccctk tksgccmtca kcayrgrtcg cta
36 <210> SEQ ID NO: 2
37 <211> LENGTH: 35
38 <212> TYPE: DNA
39 <213> ORGANISM: Artificial Sequence
41 <220> FEATURE:
42 <223> OTHER INFORMATION: primer
45 <400> SEQUENCE: 2
46 aagcttatga araaggcag scarcagagg kyrma
48 <210> SEQ ID NO: 3
49 <211> LENGTH: 33
50 <212> TYPE: DNA
51 <213> ORGANISM: Artificial Sequence
53 <220> FEATURE:
54 <223> OTHER INFORMATION: primer
57 <400> SEQUENCE: 3
58 aagcatagca ggaaggcctt gaaggccagc ctg
60 <210> SEQ ID NO: 4
61 <211> LENGTH: 33
62 <212> TYPE: DNA
63 <213> ORGANISM: Artificial Sequence
65 <220> FEATURE:
66 <223> OTHER INFORMATION: primer
69 <400> SEQUENCE: 4
70 ggcqagaaat acgcccgtaa gttctcccg gac
72 <210> SEQ ID NO: 5
73 <211> LENGTH: 30
74 <212> TYPE: DNA
75 <213> ORGANISM: Artificial Sequence

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/829,631A

DATE: 10/26/2001
TIME: 13:23:11

Input Set : A:\NIH047.1CP1C1.TXT
Output Set: N:\CRF3\10262001\I829631A.raw

77 <220> FEATURE:
 78 <223> OTHER INFORMATION: primer
 81 <400> SEQUENCE: 5
 82 tttcccaatac tactctaagg tgcatgttcc 30
 84 <210> SEQ ID NO: 6
 85 <211> LENGTH: 30
 86 <212> TYPE: DNA
 87 <213> ORGANISM: Artificial Sequence
 89 <220> FEATURE:
 90 <223> OTHER INFORMATION: primer
 93 <400> SEQUENCE: 6
 94 cacacgactt aactccatag agtcgatcg 30
 96 <210> SEQ ID NO: 7
 97 <211> LENGTH: 1914
 98 <212> TYPE: DNA
 99 <213> ORGANISM: Rat
 101 <220> FEATURE:
 102 <221> NAME/KEY: CDS
 103 <222> LOCATION: (439)...(1749)
 105 <400> SEQUENCE: 7
 106 ccaaccctt cgcgcgacac gtgggtgatct aacgtactca cacgcccacc ctttcgaag 60
 107 agactgcccc ggccggaaagg cggggatccg gtcctctgtc ccacatcccc agctgtgcc 120
 108 cttagccagg accccacccc catcttatgg catccccggt gcccctattc catccccagg 180
 109 ctctcatcca gccccaaatc aactttcatt gactcgtcac atcagtaacc ctccccaaac 240
 110 ttcttaccgg agtactccag gtggccctgc gttaggaggca cccctacaac tcctcccgat 300
 111 ctcttgaat cgctgtcga tgacctaaaga accccgtttt gccaataacta ctctaagg 360
 112 cagtttcctt tctcccttgc cttttttttt tcccttccacc ctgtacccatc tcccttccat 420
 113 gtcctcaacc cagtcccc atg gtt cca gag cca ggc cct gtc aac agt agc 471
 114 Met Val Pro Glu Pro Gly Pro Val Asn Ser Ser
 115 1 5 10
 117 acc cca gcc tgg ggt ccc ggg cca ccg cct gct ccg ggg ggc agc ggc 519
 118 Thr Pro Ala Trp Gly Pro Gly Pro Pro Ala Pro Gly Gly Ser Gly
 119 15 20 25
 121 tgg gtg gct gcc gcg ctg tgc gtg gtc atc gtg ctg aca gca gcc ggc 567
 122 Trp Val Ala Ala Leu Cys Val Val Ile Val Leu Thr Ala Ala Ala
 123 30 35 40
 125 aat tcg ctg ctg atc gtg ctc att tgc acg cag ccc gcc gtg cgc aac 615
 126 Asn Ser Leu Leu Ile Val Leu Ile Cys Thr Gln Pro Ala Val Arg Asn
 127 45 50 55
 129 acg tct aac ttc ttt ctg gtg tcg ctc ttc acg tcg gac ttg atg gtg 663
 130 Thr Ser Asn Phe Phe Leu Val Ser Leu Phe Thr Ser Asp Leu Met Val
 131 60 65 70 75
 133 ggg ttg gtg gtg atg ccc cca gcc atg ctg aac gcg ctg tat ggg cgc 711
 134 Gly Leu Val Val Met Pro Pro Ala Met Leu Asn Ala Leu Tyr Gly Arg
 135 80 85 90
 137 tgg gtg tta gct cga ggc ctc tgt ctg ctt tgg act gcc ttc gac gtg 759
 138 Trp Val Leu Ala Arg Gly Leu Cys Leu Leu Trp Thr Ala Phe Asp Val
 139 95 100 105
 141 atg tgc tgc agc gcc tcc atc ctc aac ctc tgc ctc atc agc ctg gac 807

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/829,631A

DATE: 10/26/2001
TIME: 13:23:11

Input Set : A:\NIH047.1CP1C1.TXT
Output Set: N:\CRF3\10262001\I829631A.raw

142	Met	Cys	Cys	Ser	Ala	Ser	Ile	Leu	Asn	Leu	Cys	Leu	Ile	Ser	Leu	Asp	
143		110					115						120				855
145	cgc	tac	ctg	ctc	atc	ctc	tcg	ccg	ctg	cgc	tac	aag	ctg	cgc	atg	aca	
146	Arg	Tyr	Leu	Leu	Ile	Leu	Ser	Pro	Leu	Arg	Tyr	Lys	Leu	Arg	Met	Thr	
147		125					130					135					
149	gcc	ccg	cga	gcc	ctg	gcg	ctc	atc	ctg	ggt	gcc	tgg	agc	ctc	gcg	gct	903
150	Ala	Pro	Arg	Ala	Leu	Ala	Leu	Ile	Leu	Gly	Ala	Trp	Ser	Leu	Ala	Ala	
151	140						145					150			155		
153	ctt	gcc	tcc	tcc	cta	ccc	ctc	ttg	ctg	ggc	tgg	cac	caa	ctg	qgc	aaa	951
154	Leu	Ala	Ser	Phe	Leu	Pro	Leu	Leu	Gly	Trp	His	Glu	Leu	Gly	Lys		
155		160					165					170					
157	gct	cga	aca	cct	gcc	cct	ggc	cag	tgc	cgc	cta	ttg	gcc	agc	ctg	cct	999
158	Ala	Arg	Thr	Pro	Ala	Pro	Gly	Gln	Cys	Arg	Leu	Leu	Ala	Ser	Leu	Pro	
159		175					180					185					
161	ttt	gtc	ctc	gtg	gcc	tcc	ggc	gtc	acc	ttt	ttc	ctg	cct	tcg	ggt	gcc	1047
162	Phe	Val	Leu	Val	Ala	Ser	Gly	Val	Thr	Phe	Phe	Leu	Pro	Ser	Gly	Ala	
163		190					195					200					
165	atc	tgc	tcc	acc	tac	tgc	agg	atc	ctt	ctg	gct	gcc	cgc	aag	cag	gct	1095
166	Ile	Cys	Phe	Thr	Tyr	Cys	Arg	Ile	Leu	Leu	Ala	Ala	Arg	Lys	Gln	Ala	
167		205					210					215					
169	gtg	caa	gtg	gcc	tcg	ctc	acc	acg	ggc	acg	gct	ggc	cag	gcc	ttg	gaa	1143
170	Val	Gln	Val	Ala	Ser	Leu	Thr	Thr	Gly	Thr	Ala	Gly	Gln	Ala	Leu	Glu	
171	220						225					230			235		
173	acc	ttg	cag	gtg	ccc	agg	aca	cca	cgc	cca	ggg	atg	gag	tcc	gct	gac	1191
174	Thr	Leu	Gln	Val	Pro	Arg	Thr	Pro	Arg	Pro	Gly	Met	Glu	Ser	Ala	Asp	
175		240					245					250					
177	agt	agg	cgt	ctg	gcc	acc	aag	cat	agc	agg	aag	gcc	ttg	aag	gcc	agc	1239
178	Ser	Arg	Arg	Leu	Ala	Thr	Lys	His	Ser	Arg	Lys	Ala	Leu	Lys	Ala	Ser	
179		255					260					265					
181	ctg	acc	ctg	ggc	atc	ctg	ctg	gga	atg	ttc	ttt	gtc	acc	tgg	ctg	ccc	1287
182	Leu	Thr	Leu	Gly	Ile	Leu	Leu	Gly	Met	Phe	Phe	Val	Thr	Trp	Leu	Pro	
183		270					275					280					
185	ttc	ttt	gtg	gcc	aac	ata	gct	cag	gcc	gtg	tgt	gac	tgc	atc	tcc	cca	1335
186	Phe	Phe	Val	Ala	Asn	Ile	Ala	Gln	Ala	Val	Cys	Asp	Cys	Ile	Ser	Pro	
187		285					290					295					
189	ggc	ctc	tcc	gat	gtc	ctc	aca	tgg	ctg	ggg	tac	tgt	aat	agc	acc	atg	1383
190	Gly	Leu	Phe	Asp	Val	Leu	Thr	Trp	Leu	Gly	Tyr	Cys	Asn	Ser	Thr	Met	
191	300						305					310			315		
193	aac	cct	atc	atc	tac	ccg	ctc	ttt	atg	cgg	gac	ttc	aag	agg	gcc	ctg	1431
194	Asn	Pro	Ile	Ile	Tyr	Pro	Leu	Phe	Met	Arg	Asp	Phe	Lys	Arg	Ala	Leu	
195		320					325					330					
197	ggc	agg	ttc	ctg	cat	gct	tcc	act	gtc	ccc	cg	agc	acc	ggc	cag	ccc	1479
198	Gly	Arg	Phe	Leu	His	Ala	Ser	Thr	Val	Pro	Arg	Ser	Thr	Gly	Gln	Pro	
199		335					340					345					
201	tgc	ctc	ccc	ctc	cat	gtg	gac	ctc	tca	cag	cg	tgc	cag	acc	agg	cct	1527
202	Cys	Leu	Pro	Leu	His	Val	Asp	Leu	Ser	Gln	Arg	Cys	Gln	Thr	Arg	Pro	
203		350					355					360					
205	cag	ctg	cag	cag	gtg	ctc	gct	ctg	cct	ctg	ccg	cca	aac	tca	gat	tca	1575
206	Gln	Leu	Gln	Gln	Val	Leu	Ala	Leu	Pro	Leu	Pro	Pro	Asn	Ser	Asp	Ser	

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/829,631A

DATE: 10/26/2001
TIME: 13:23:11

Input Set : A:\NIH047.1CP1C1.TXT
Output Set: N:\CRF3\10262001\I829631A.raw

207	365	370	375	
209	gac tcc gct tca ggg ggc acc tcg ggc ctg cag ctc aca gcc cag ctt			1623
210	Asp Ser Ala Ser Gly Gly Thr Ser Gly Leu Gln Leu Thr Ala Gln Leu			
211	380	385	390	395
213	ctg ctg cct gga gag gcc aca cgg gac ccc ccg cca ccc acc agg gcc			1671
214	Leu Leu Pro Gly Glu Ala Thr Arg Asp Pro Pro Pro Pro Thr Arg Ala			
215	400	405	410	
217	acc act gtg gtc aac ttc ttt gtc aca gac tct gtg gag cct gag ata			1719
218	Thr Thr Val Val Asn Phe Phe Val Thr Asp Ser Val Glu Pro Glu Ile			
219	415	420	425	
221	cgg ccg cat cca ctc agt tcc ccc gtg aac tgaccaggc aagagctggc			1769
222	Arg Pro His Pro Leu Ser Ser Pro Val Asn			
223	430	435		
225	cattggaggc cacatccccg gagctctcaag cccactctcc ctgagactag gaggtggtag			1829
226	gtctccttag agtgtgctga attgaggtat ctcagctgc ccatcttctg ctgcagctcc			1889
227	ttgactgagg ggtagtcaga cacat			1914
229	<210> SEQ ID NO: 8			
230	<211> LENGTH: 437			
231	<212> TYPE: PRT			
232	<213> ORGANISM: Rat			
234	<400> SEQUENCE: 8			
235	Met Val Pro Glu Pro Gly Pro Val Asn Ser Ser Thr Pro Ala Trp Gly			
236	1	5	10	15
237	Pro Gly Pro Pro Pro Ala Pro Gly Gly Ser Gly Trp Val Ala Ala Ala			
238	20	25	30	
239	Leu Cys Val Val Ile Val Leu Thr Ala Ala Ala Asn Ser Leu Leu Ile			
240	35	40	45	
241	Val Leu Ile Cys Thr Gln Pro Ala Val Arg Asn Thr Ser Asn Phe Phe			
242	50	55	60	
243	Leu Val Ser Leu Phe Thr Ser Asp Leu Met Val Gly Leu Val Val Met			
244	65	70	75	80
245	Pro Pro Ala Met Leu Asn Ala Leu Tyr Gly Arg Trp Val Leu Ala Arg			
246	85	90	95	
247	Gly Leu Cys Leu Leu Trp Thr Ala Phe Asp Val Met Cys Cys Ser Ala			
248	100	105	110	
249	Ser Ile Leu Asn Leu Cys Leu Ile Ser Leu Asp Arg Tyr Leu Leu Ile			
250	115	120	125	
251	Leu Ser Pro Leu Arg Tyr Lys Leu Arg Met Thr Ala Pro Arg Ala Leu			
252	130	135	140	
253	Ala Leu Ile Leu Gly Ala Trp Ser Leu Ala Ala Leu Ala Ser Phe Leu			
254	145	150	155	160
255	Pro Leu Leu Leu Gly Trp His Glu Leu Gly Lys Ala Arg Thr Pro Ala			
256	165	170	175	
257	Pro Gly Gln Cys Arg Leu Leu Ala Ser Leu Pro Phe Val Leu Val Ala			
258	180	185	190	
259	Ser Gly Val Thr Phe Phe Leu Pro Ser Gly Ala Ile Cys Phe Thr Tyr			
260	195	200	205	
261	Cys Arg Ile Leu Leu Ala Ala Arg Lys Gln Ala Val Gln Val Ala Ser			
262	210	215	220	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/829,631A

DATE: 10/26/2001

TIME: 13:23:11

Input Set : A:\NIH047.1CP1C1.TXT

Output Set: N:\CRF3\10262001\I829631A.raw

263 Leu Thr Thr Gly Thr Ala Gly Gln Ala Leu Glu Thr Leu Gln Val Pro
 264 225 230 235 240
 265 Arg Thr Pro Arg Pro Gly Met Glu Ser Ala Asp Ser Arg Arg Leu Ala
 266 245 250 255
 267 Thr Lys His Ser Arg Lys Ala Leu Lys Ala Ser Leu Thr Leu Gly Ile
 268 260 265 270
 269 Leu Leu Gly Met Phe Phe Val Thr Trp Leu Pro Phe Phe Val Ala Asn
 270 275 280 285
 271 Ile Ala Gln Ala Val Cys Asp Cys Ile Ser Pro Gly Leu Phe Asp Val
 272 290 295 300
 273 Leu Thr Trp Leu Gly Tyr Cys Asn Ser Thr Met Asn Pro Ile Ile Tyr
 274 305 310 315 320
 275 Pro Leu Phe Met Arg Asp Phe Lys Arg Ala Leu Gly Arg Phe Leu His
 276 325 330 335
 277 Ala Ser Thr Val Pro Arg Ser Thr Gly Gln Pro Cys Leu Pro Leu His
 278 340 345 350
 279 Val Asp Leu Ser Gln Arg Cys Gln Thr Arg Pro Gln Leu Gln Gln Val
 280 355 360 365
 281 Leu Ala Leu Pro Leu Pro Pro Asn Ser Asp Ser Ala Ser Gly
 282 370 375 380
 283 Gly Thr Ser Gly Leu Gln Leu Thr Ala Gln Leu Leu Leu Pro Gly Glu
 284 385 390 395 400
 285 Ala Thr Arg Asp Pro Pro Pro Pro Thr Arg Ala Thr Thr Val Val Asn
 286 405 410 415
 287 Phe Phe Val Thr Asp Ser Val Glu Pro Glu Ile Arg Pro His Pro Leu
 288 420 425 430
 289 Ser Ser Pro Val Asn
 290 435
 293 <210> SEQ ID NO: 9
 294 <211> LENGTH: 2108
 295 <212> TYPE: DNA
 296 <213> ORGANISM: Rat
 298 <220> FEATURE:
 299 <221> NAME/KEY: CDS
 300 <222> LOCATION: (439)...(1311)
 302 <221> NAME/KEY: intron
 303 <222> LOCATION: (1312)...(1505)
 305 <221> NAME/KEY: CDS
 306 <222> LOCATION: (1506)...(1943)
 308 <400> SEQUENCE: 9
 309 ccaaccccca cgcgcgacac gtgggtatct aacgtactca cacgcccacc cttctcgaaag 60
 310 agactgcccc ggccggaagg cgggagttcg gctcctgcgc ccacatcccc agctgtgccc 120
 311 ctagccagga accccacccc catcttatgg catccccggt gcccctattc catccccagg 180
 312 ctctcatcca gccccaaagct aactttcatt gactcgtcac atcagtaacc ctccccaaac 240
 313 ttcttacccg agtactccag gtggccctgc gttaggaggca cccctacaac tcctcccgat 300
 314 ctcttgaat cgctgctcga tgacctaaga accccgtttt gccaataacta ctctaaggtg 360
 315 caqcttcatt ttccttcatt tgccttcacc ctgtacactgc agtcaccata tccctgtctg 420
 316 gtctcaacc cagtcccc atg gtt cca gag cca ggc cct gtc aac agt agc 471
 317 Met Val Pro Glu Pro Gly Pro Val Asn Ser Ser

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/829,631A

DATE: 10/26/2001

TIME: 13:23:12

Input Set : A:\NIH047.1CP1C1.TXT

Output Set: N:\CRF3\10262001\I829631A.raw

L:532 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:533 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:633 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:658 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:708 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13